

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd

OM protein - protein search, using sw model

Run on: January 16, 2003, 16:34:37 : Search time 24 7857 seconds
(without alignments)
28.011 Million cell updates/sec

Title: us-09-856-070-26

Perfect score: 28

Sequence: 1 query 5

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 908470 seqs, 13250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 99

Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_101002.*

- 1: /SID22/qcdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID22/qcdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID22/qcdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID22/qcdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID22/qcdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID22/qcdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID22/qcdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID22/qcdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID22/qcdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID22/qcdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID22/qcdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID22/qcdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID22/qcdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID22/qcdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID22/qcdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID22/qcdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID22/qcdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID22/qcdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID22/qcdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID22/qcdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID22/qcdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID22/qcdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID22/qcdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	5	AA82201	Human heprecceptor
2	28	100.0	11	AA82203	Human heprecceptor
3	28	100.0	11	AA82203	Human heprecceptor
4	28	100.0	14	AA82203	Human heprecceptor
5	28	100.0	27	AA82203	Antennapedia inter
6	28	100.0	34	AA82203	Human heprecceptor
7	28	100.0	47	AA82203	Human bone marrow
8	28	100.0	52	AA82203	Novel human secret
9	28	100.0	137	AA82203	Streptococcus pneu
10	28	100.0	159	AA82203	Human PPO polyref

11	28	100.0	168	21	AA70434	Haematobia irritan
12	28	100.0	175	21	AA70435	Haematobia irritan
13	28	100.0	252	22	AA82203	Human protein sequ
14	28	100.0	252	22	AA82203	Human protein sequ
15	28	100.0	313	22	AA82203	Novel human diagno
16	28	100.0	348	22	AA82203	Putative f. abysssi
17	28	100.0	400	22	AA82203	Drosophila melanog
18	28	100.0	406	22	AA82203	Human protein sequ
19	28	100.0	436	22	AA82203	Human colon cancer
20	28	100.0	478	22	AA82203	Novel human enzyme
21	28	100.0	490	22	AA82203	Novel human enzyme
22	28	100.0	516	22	AA82203	C. pneumoniae prot
23	28	100.0	537	22	AA82203	Human protein sequ
24	28	100.0	568	23	AA82203	Arabidopsis thausc
25	28	100.0	586	20	AA82203	Amino acid sequen
26	28	100.0	622	22	AA82203	Novel human secret
27	28	100.0	635	21	AA82203	Human colon cancer
28	28	100.0	824	22	AA82203	Human protein sequ
29	28	100.0	937	22	AA82203	Novel human diagno
30	28	100.0	1024	23	AA82203	LDL receptor bindi
31	28	100.0	1342	20	AA82203	Human IKAP protein
32	28	100.0	1332	22	AA82203	Human protein sequ
33	28	100.0	1332	22	AA82203	Human IKAP. Homo
34	25	89.3	9	18	AA82203	Human IKAP. Homo
35	25	89.3	14	20	AA82203	Band III-GST fusio
36	25	89.3	27	19	AA82203	Deletion sequence
37	25	89.3	36	19	AA82203	Part of the gp75 p
38	25	89.3	36	19	AA82203	1514A substitution
39	25	89.3	36	22	AA82203	Peptide #6075 enco
40	25	89.3	36	22	AA82203	Protein #634 enco
41	25	89.3	36	22	AA82203	Human brain extrac
42	25	89.3	36	22	AA82203	Human bone marrow
43	25	89.3	36	22	AA82203	Peptide #5724 enco
44	25	89.3	36	22	AA82203	Peptide #6057 enco
45	25	89.3	36	23	AA82203	Human peptide enco

ALIGNMENTS

RESULT 1	AA822041	standard, peptide, 5 AA.
ID	AA822041	standard, peptide, 5 AA.
AC	AA822041	
XX	13-JUN-2001	(first entry)
XX	Human heprecceptor domain A binding peptide- Eup-2832.	
XX	Human heprecceptor, cytosolic; anti-HIV; antibiotic;	
XX	Human heprecceptor, cytosolic; anti-HIV; antibiotic;	
XX	HIV related dementia	
XX	Homo sapiens.	
XX	Key	location/Qualifiers
XX	Modified-site	3
XX	/note	"Optionally phosphorylated"
XX	GB2354241-A.	
XX	21-MAR-2001.	
XX	17-SEP-1999.	99GB 0021881.
XX	17-SEP-1999.	99GB 0021881.
XX	(HOLM/) HOLMS R.D.	
XX	Holms RD:	
XX	WET, 2001 253287/31.	


```

RESULT 4
AAB82037
ID AAB82037 standard; peptide: 13 AA.
AC AAB82037;
XX
XX 13-JUN-2001 (first entry)
XX Human hepreceptor domain A binding peptide 9upe2042.
XX
XX Human; hepreceptor; cytostatic; anti-HIV; antibiotic;
KW neotropic; immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified site 11
XX /note- "Optionally phosphorylated"
XX
XX GB2354241-A.
XX
XX 21-MAR-2001.
XX
XX 17-SEP-1999; 9qGR-0021881.
XX
XX 17-SEP-1999; 9qGR-0021881.
XX (HOLM/) HOLMS R D.
XX Holms RD;
XX
XX WPI: 2001-293287/41.
XX
XX Novel regulatory or scaffolding peptides of ezrin that binds to
XX hepreceptor, useful for inducing immune response for treating
XX infectious diseases and cancer.
XX
XX Claim 22: Page 36 42pp, English
XX
XX The hepreceptor is a novel active site in human ezrin. Ezrin regulates
XX the structure of the cortical cytoskeleton to control cell surface
XX topography. The present invention relates to peptides (See AAB82041 to
XX AAB82041) that bind to hepreceptor with greater affinity than HEPI (see
XX AAB82046) the hepreceptor binding peptides are useful for inducing
XX immune response, and for treating infectious diseases, cancer and
XX HIV-related dementia. The present peptide binds to domain A of the
XX hepreceptor (AAB82019).
XX
XX Sequence 13 AA;
XX
XX Query Match 100.0%; Score 28; DB 22; Length 13;
XX Best Local Similarity 100.0%; Pred. No. 28;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 QDYEE 5
XX
XX 9 QDYEE 13
XX
XX RESULT 5
XX AAY27444
XX ID AAY27444 standard; peptide: 27 AA.
XX
XX AC AAY27444;
XX
XX 26-NOV-1999 (first entry)
XX
XX Antennapedia internalization sequence in tandem with ezrin fragment.
XX
XX Pharmaceutical; ezrin; mutant; tumor; antennapedia internalization;
KW metastasis; human.

```

```

XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 1
XX /note- "biotinylated"
XX
XX Modified-site 22
XX /note- "Optionally phosphorylated"
XX
XX W09947150 A2.
XX
XX 23-SEP-1999.
XX
XX 18-MAR-1999; 99WO FP02054.
XX
XX 18-MAR-1998; 98US-0040725.
XX
XX (CURT-) INST CURIF.
XX (CNRS ) CNRS CENT NA; KCH SCI.
XX
XX Arpin M, Crepaldi T, Gautreau A, Louvard D;
XX WPI: 1999-561851/47.
XX
XX New composition for prevention and treatment of tumors and metastasis
XX
XX Example 5; Page 14: 31pp; English.
XX
XX The invention provides a pharmaceutical composition containing ezrin
XX protein, RNA or DNA mutated on tyrosine 353, or a functional fragment
XX or derivative of the ezrin mutant. The new composition is useful for
XX prevention and/or treatment of tumors, and especially metastasis. The
XX present sequence represents an antennapedia internalization sequence in
XX tandem with an ezrin fragment (residues 348-358). This is used in
XX experiments of p85 interaction with phosphorylated ezrin peptides.
XX
XX Sequence 27 AA;
XX
XX Query Match 100.0%; Score 28; DB 25; Length 27;
XX Best Local Similarity 100.0%; Pred. No. 57;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 QDYEE 5
XX
XX 20 QDYEE 24
XX
XX RESULT 6
XX AAB82020
XX ID AAB82020 standard; peptide: 34 AA.
XX
XX AC AAB82020;
XX
XX 13-JUN-2001 (first entry)
XX
XX Human hepreceptor domain B.
XX
XX Human; hepreceptor domain B; cytostatic; anti-HIV; antibiotic;
KW neotropic; immune response inducer; ezrin; infectious diseases; cancer;
KW HIV-related dementia.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Modified-site 14
XX /note- "Optionally phosphorylated"
XX
XX GB2354241-A.
XX
XX 21-MAR-2001.
XX
XX 17-SEP-1999; 9qGR-0021881

```

```

XX 17-SEP-1999; 99GB-0021881.
XX (HOLM/Z) HOLMS R D.
XX Holms RD;
XX WPI: 2001-293287/41.
XX Novel regulatory or unfolding peptides of ezrin that binds to
XX heptareceptor, useful for inducing immune response for treating
XX infectious diseases and cancer.
XX Claim 5; Page 36; 42pp; English.
XX The present sequence is domain H of human heptareceptor of human ezrin. The
XX heptareceptor is a novel active site in human ezrin. Ezrin regulates the
XX structure of the cortical cytoskeleton to control cell surface
XX topography. The present invention relates to peptides (see AAB82021 to
XX AAB82041) that bind to heptareceptor with greater affinity than HEPL (see
XX AAB82046). The heptareceptor binding peptides are useful for inducing
XX immune response, and for treating infectious diseases, cancer and
XX HIV-related dementia. The present sequence assembles into two
XX anti parallel helices with heptareceptor domain A (see AAB82019)
XX Sequence 34 AA;
XX
XX Query Match 100.0%; Score 28; DB 22; Length 34;
XX Best Local Similarity 100.0%; Pred. No. 71;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 QYEE 5
XX I I I I I
XX DB 12 QYEE 16
XX
XX RESULT 7
XX AAM78270
XX ID AAM78270 standard; Protein; 37 AA.
XX AC
XX AAM78270;
XX DB 06-NOV 2001 (first entry)
XX
XX Human bone marrow expressed proto encoded protein SEQ ID NO: 38576.
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX Homo sapiens.
XX
XX W0200157276-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-050964B.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0287456.
XX 30-JUN-2000; 2000US-0638408.
XX 03-AUG-2000; 2000US-0642364.
XX 21-SEP-2000; 2000US-0234687.
XX 27-SEP-2000; 2000US-0236359.
XX 04-OCT-2000; 2000CA-0021263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow.

```

```

XX Example 4: SEQ ID NO: 38576; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is a
XX protein encoded by one of the probes of the invention.
XX Sequence 37 AA;
XX
XX Query Match 100.0%; Score 28; DB 22; Length 37;
XX Best Local Similarity 100.0%; Pred. No. 77;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 QYEE 5
XX I I I I I
XX DB 19 QYEE 23
XX
XX RESULT 8
XX AAU33060
XX ID AAU33060 standard; Protein; 52 AA.
XX AC
XX AAU33060;
XX DB 18-DEC 2001 (first entry)
XX
XX Novel human secreted protein #3551.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
XX stem cell proliferation; haematopoiesis; nerve tissue regeneration;
XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.
XX Homo sapiens.
XX
XX W0200179449-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-0508656.
XX 18-APR-2000; 2000US-0552929.
XX 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
XX PA
XX Tang YT, Liu C, Dimaheac RT;
XX WPI: 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
XX vaccination, testing and therapy.
XX Claim 20, Page 702, 765pp, English.
XX
XX The invention relates to novel human secreted polypeptides, the
XX polypeptides and antibodies to the polypeptides are useful for
XX determining the presence of or predisposition to a disease associated
XX with altered levels of polypeptide. The polypeptides are also useful for
XX identifying agents (agonists and antagonists) that bind to them. Cells
XX expressing the proteins are useful for identifying a therapeutic agent
XX for use in treatment of a pathology related to aberrant expression of
XX physiological interactions of the polypeptide. Vectors comprising
XX the nucleic acids encoding the polypeptides and cells genetically
XX engineered to express them are also useful for producing the proteins.
XX The proteins are useful in genetic vaccination, testing and
XX therapy, and can be used as nutritional supplements. They may be used to
XX increase stem cell proliferation; to regulate haematopoiesis; and in
XX bone, cartilage, tendon and/or nerve tissue growth or regeneration;
XX immune suppression and/or stimulation; as anti-inflammatory agents; and
XX in treatment of leukaemias. AAU29510-AAU33304 represent the amino acid

```

CC sequences of novel human secreted proteins of the invention.

XX Sequence 52 AA;

SO Query Match

Best Local Similarity 100.0%; Score 28; DB 22; Length 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QDYEE 5

DB 20 QDYEE 24

RESULT 9

AAW62697 ID AAW62697 standard; Protein: 137 AA.

XX AC AAW62697;

XX DT 09-NOV-1998 (first entry)

XX DE Streptococcus pneumoniae Polypeptide

XX KW Polypeptide ORF open reading frame; infection; bacterial; streptococcal, bacteremia, diagnosis, prophylaxis.

XX OS Streptococcus pneumoniae.

XX PN W09823631-A1

XX PD 04-FUN-1998

XX PP 24-NOV-1997; 97W0-US21976.

XX PR 27-NOV-1996; 96NS-0031879.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PI Black MT, Hodgson JE, Knowles DJ, Lenette MA, Williams RJ, Reid RH, Zarfos PN;

XX PP WPI: 1998-322654/28

XX PT Streptococcus pneumoniae polypeptides - useful for developing products for diagnosis, prevention and treatment of infections e.g. pneumonia, bacteremia, meningitis or endocarditis

XX PS Claim 5; Page 32; 18pp; English

XX CC The sequence is that of a Streptococcal polypeptide

XX CC The polypeptide can potentially be used for the diagnosis and prevention of bacterial infections, especially septicemia.

XX CC It may be used for the treatment of diseases such as otitis media, conjunctivitis, pneumonia, bacteremia, meningitis, sinusitis, pleural empyema, endocarditis or infection of the cerebrospinal fluid

XX SO Sequence 137 AA;

Query Match

Best Local Similarity 100.0%; Score 28; DB 19; Length 137;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QDYEE 5

DB 38 QDYEE 42

RESULT 10

AAU29293

ID AAU29293 standard; Protein: 159 AA.

XX AC AAU29293;

XX DT 18-DEC-2001 (first entry)

XX Human PRO polypeptide sequence #270.

XX PRO polypeptide, mammal, tumour, cancer, human, cattle, horse, sheep; dog, cat, pig, goat, rabbit, tumour necrosis factor alpha; TNF-alpha; blood; chondrocyte cell; cell proliferation; cell differentiation; colon; adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX Homo sapiens.

XX W0200168848-A2.

XX 20-SEP-2001.

XX 28 FTH 2001; 2001W0-US05520.

XX 01-MAR-2000; 2000W0-US05601.

XX 02-MAR-2000; 2000W0-US05841.

XX 03-MAR-2000; 2000US-187202P.

XX 05-MAR-2000; 2000US-186968P.

XX 14-MAR-2000; 2000US-189420P.

XX 14-MAR-2000; 2000US-189428P.

XX 15-MAR-2000; 2000W0-US06084.

XX 21-MAR-2000; 2000US-190828P.

XX 21-MAR-2000; 2000US-191007P.

XX 21-MAR-2000; 2000US-191048P.

XX 21-MAR-2000; 2000US-191314P.

XX 28-MAR-2000; 2000US-192655P.

XX 29-MAR-2000; 2000US-193032P.

XX 30-MAR-2000; 2000US-193053P.

XX 04-APR-2000; 2000W0-US08439.

XX 04-APR-2000; 2000US-194449P.

XX 11-APR-2000; 2000US-195975P.

XX 11-APR-2000; 2000US-196000P.

XX 11-APR-2000; 2000US-196187P.

XX 11-APR-2000; 2000US-196690P.

XX 17-APR-2000; 2000US-196820P.

XX 18-APR-2000; 2000US-198121P.

XX 18-APR-2000; 2000US-198585P.

XX 25-APR-2000; 2000US-199497P.

XX 25-APR-2000; 2000US-199550P.

XX 25-APR-2000; 2000US-199654P.

XX 03-MAY-2000; 2000US-201516P.

XX 17-MAY-2000; 2000W0-US13705.

XX 22-MAY-2000; 2000W0-US14042.

XX 30-MAY-2000; 2000W0-US14941.

XX 02-JUN-2000; 2000W0-US15264.

XX 06-JUN-2000; 2000US-203832P.

XX 28-JUL-2000; 2000W0-US20710.

XX 22-AUG-2000; 2000US-064848.

XX 24-AUG-2000; 2003W0-US23328.

XX 08-NOV-2000; 2000W0-US30952.

XX 01-DEC-2000; 2003W0-US32678.

XX 20-DEC-2000; 2000W0-US34956.

(GETH) GENENTECH INC.

Baker KP, Chen J, Desoyers L, Gokkard A, Gadowski RJ, Gurney AL,

Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;

WPI: 2001-602746/88.

N-PSDH; AAS46194.

Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

Claim 11; Fig 540; 774pp; English.

Sequences AA229024 AAU29298 represent PRO polypeptides of the invention.

CC the pro polypeptides and their associated nucleic acids can be used to
 CC detect the presence of a tumour in a mammal by comparing the level of
 CC expression of a pro polypeptide in a test sample of cells from the animal
 CC and a control sample of normal cells, whereby a higher level of
 CC expression in the test sample indicates the presence of a tumour in the
 CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
 CC and rabbits but are preferably human. The polypeptides can be used to
 CC stimulate tumour necrosis factor (TNF) alpha release from human blood
 CC when contacted with it. A specific polypeptide can be used to stimulate
 CC the proliferation or differentiation of chondrocyte cells. The pro
 CC proteins can be used to determine the presence of tumours and also
 CC susceptibility to tumour development, particularly adrenal, lung, colon,
 CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
 CC subjects. The oligonucleotide probes specific for the pro nucleic acids
 CC can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 159 AA;
 Query Match 100.0%; Score 28; DB 22; Length 159;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYEE 5

DB 22 QYEE 26

RESULT 11

AAAY70435
 ID AAY70435 standard; Protein; 168 AA.

AC AAY70435;
 DT 21-JUN-2000 (first entry)

DE Haematobia irritans partial thrombostatin protein from clone PH8.

XX Thrombostatin; horn fly; thrombin; antithrombin; thrombolytic; clone PH8;
 KW coagulation; veterinary vaccine; treatment; haematophagy; prophylaxis;
 KW thromboembolism; thrombosis; acute shock therapy; coagulopathies;
 KW haemodialysis; haemoseparation; extracorporeal blood circulation;
 KW anticoagulant.

OS Haematobia irritans.

XX Key Location/Qualifiers

FT Active site 88..168

FI /label= Active_thrombostatin

XX WC200011172-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18888.

XX 20-AUG-1998; 98US-0097227.

XX (UYAU) UNIV AUBURN.

XX Cupp EW, Cupp MS;

XX WPI: 2000-246563/21

XX N-PSDB; AAZ51575.

XX Proteins with antithrombin activity useful as veterinary vaccines for

XX prevention and treatment of hematophagy in cattle and thrombosis in

XX mammals are isolated from the salivary glands of horn fly.

XX Claim 4b; Page 53; 58pp; English.

XX The present amino acid sequence is the partial protein, encoded by clone

XX PH8, having antithrombin activity, designated as thrombostatin. It is

XX isolated from the salivary glands of Haematobia irritans (horn fly).

CC that contains an inhibitor of thrombin. The protein prevents coagulation
 CC of blood by inhibiting the activity of thrombin (factor II). It comprises
 CC 21% of aspartic acid and glutamic acid residues. This sequence has
 CC thrombolytic activity. Thrombostatin sequences are useful as veterinary
 CC vaccines for treating haematophagy (blood-feeding) in cattle and
 CC thrombosis in mammals. It is useful for therapy and prophylaxis of
 CC thrombosis and thromboembolisms in humans, including prophylaxis of
 CC post-operative thrombosis, acute shock therapy, therapy for consumption
 CC of coagulopathies, in haemodialyses, haemoseparations and extracorporeal
 CC blood circulation. It can also be used as an anticoagulant in blood.

XX Sequence 168 AA;

Query Match 100.0%; Score 28; DB 21; Length 168;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QYEE 5

DB 76 QYEE 80

RESULT 12

AAAY70435

ID AAY70435 standard; Protein; 175 AA.

AC AAY70435;

XX 21-JUN-2000 (first entry)

DE Haematobia irritans thrombostatin full-length protein.

XX Thrombostatin; horn fly; thrombin; antithrombin; thrombolytic;
 KW coagulation; veterinary vaccine; treatment; haematophagy; prophylaxis;
 KW thromboembolism; thrombosis; acute shock therapy; coagulopathies;
 KW haemodialysis; haemoseparation; extracorporeal blood circulation;
 KW anticoagulant.

OS Haematobia irritans.

XX Key Location/Qualifiers

FT Active site 95..175

FI /label= Active_thrombostatin

XX WC200011172-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18888.

XX 20-AUG-1998; 98US-0097227.

XX (UYAU) UNIV AUBURN.

XX Cupp EW, Cupp MS;

XX WPI: 2000-246563/21.

XX N-PSDB; AAZ51576.

XX Proteins with antithrombin activity useful as veterinary vaccines for

XX prevention and treatment of hematophagy in cattle and thrombosis in

XX mammals are isolated from the salivary glands of horn fly.

XX Claim 4b; Page 55; 58pp; English.

XX The present amino acid sequence is the full length protein having
 CC antithrombin activity, designated as thrombostatin. It is isolated from
 CC the salivary glands of Haematobia irritans (horn fly), that contains
 CC an inhibitor of thrombin. The protein prevents coagulation of blood by
 CC inhibiting the activity of thrombin (factor II). It comprises 21% of
 CC aspartic acid and glutamic acid residues. This sequence has thrombolytic
 CC activity. Thrombostatin sequences are useful as veterinary vaccines for
 CC treating haematophagy (blood feeding) in cattle and thrombosis in

CC mammals. It is useful for therapy and prophylaxis of thrombosis and
 CC thromboembolisms in humans, including prophylaxis of post-operative
 CC thrombosis, acute shock therapy, therapy for consumption of
 CC coagulopathies, in haemodialyses, haemoseparations and extracorporeal
 CC blood circulation. It can also be used as an anticoagulant in blood.
 XX
 SQ Sequence 175 AA;
 Query Match 100.0%; Score 28; DR 21; Length 175;
 Best Local Similarity 100.0%; Pred. No. 3 4e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDYEE 5
 DB 83 QDYEE 87
 RESULT 13
 AAB93862
 ID AAB93862 standard; Protein; 252 AA.
 AC AAB93862;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:13733.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PP 20-JUL-1999; 99JP-0248036.
 PP 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 WPI; 2001-318749/34.
 XX
 PI primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 13733; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAB93862 is AAB93828 and
 CC AAB93633 to AAB93742 represent human cDNA sequences; AAB92446 to
 CC AAB92699 represent human amino acid sequences, and AAB93629 to AAB93632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 SQ Sequence 252 AA;
 Query Match 100.0%; Score 28; DR 22; Length 252;
 Best Local Similarity 100.0%; Pred. No. 4 9e-02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 QDYEE 5
 DB 230 QDYEE 234
 RESULT 14
 AAB94845
 ID AAB94845 standard; Protein; 252 AA.
 AC AAB94845;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:16022.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PP 29-JUL-1999; 99JP-0248036.
 PP 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 WPI; 2001-318749/34.
 XX
 PI primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 PS Claim 8; SEQ ID 16022; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification, where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAG3166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX
 SQ Sequence 252 AA;

Query Match 100.0%; Score 28; DB 22; Length 252;
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDYEE 5
 |||||
 Db 230 QDYEE 234

RESULT 15

ABG17263
 ID ABG17263 standard; Protein; 313 AA.

AC AHC17263;

XX 18-FEB-2002 (first entry)

XX Novel human diagnostic protein #17254.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-0508631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSK-) HYSEK INC.

XX Drmanac RT, Liu C, Tanq YT;

XX WPI; 2001-639362/73.

XX N-PSDH; AAS81450.

XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID No 47622; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. AAG30010 AAG30377 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at http://wipo.int/pub/published_pct_sequences.

XX SQ Sequence 313 AA;

Query Match 100.0%; Score 28; DB 22; Length 313;
 Best Local Similarity 100.0%; Pred. No. 6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QDYEE 5
 |||||
 Db 220 QDYEE 224

Search completed: January 16, 2003, 16:49:17
 Job time : 24.7857 secs